Reprinted from VIROLOGY, Volume 12, No. 4, December 1960

Comparison of the Tryptic Peptides of Chemically Induced and Spontaneous Mutants of Tobacco Mosaic Virus

The genetic information of tobacco mosaic virus (TMV) resides in the sequence of the bases within its ribonucleic acid (RNA). The infecting TMV-RNA induces the cell to produce new virus particles consisting of RNA and a specific protein. There must therefore exist a structural relation between the RNA and the corresponding protein. One way to study this relation is to change the sequence of the RNA bases and to study the effect on the corresponding protein.

Treatment of TMV-RNA with nitrous acid converts certain bases into others, e.g., uracil to cytosin (*1*); a change of one nucleotide is sufficient to produce a mutation (*2*). By variation of the experimental conditions it is possible to vary the statistically defined average number of nucleotide changes within the RNA. Investigation of the protein structure of these chemically induced mutants promises to give information on the coding problem between nucleic acids and the corresponding proteins. The results of a comparison of the tryptic peptides of 26 chemically

induced and spontaneous TMV mutants will be described here. Part of this work has been presented before the National Academy of Sciences, Washington, D. C. (*3*). Comparisons of the amino acid composition (*4*) and of the composition of the tryptic peptides (*5*) of the normal TMV strain and a mutant induced by nitrous acid treatment have been published.

To solve the coding problem, it is essential to have a method capable of detecting with certainty a change in only one out of the 157 amino acids making up the TMV protein. Therefore the purified TMV was split into RNA and protein and (besides determining the amino acid composition of the virus protein in an automatic amino acid analyzer) the protein was digested with trypsin; the resulting 12 tryptic peptides were isolated and purified by column and paper chromatography and paper electrophoresis. Their amino acid composition was analyzed in amino acid analyzers according to Spackman *et al.* (*6*). The methods were essentially the same as those described previously in detail (*5*, *7*).

Figure 1 gives a scheme showing derivation of the mutants whose tryptic peptides have been studied. Those mutants designated by "Ni" were produced by nitrous acid treatment; all others were spontaneous mutants. The isolation of the spontaneous mutants A 7 to E 83 was done in such a way as to reduce to a minimum the probability of getting more than one mutation step (*8*). This principle was not stressed as much in isolating the spontaneous mutants designated by Latin names [*flavum*

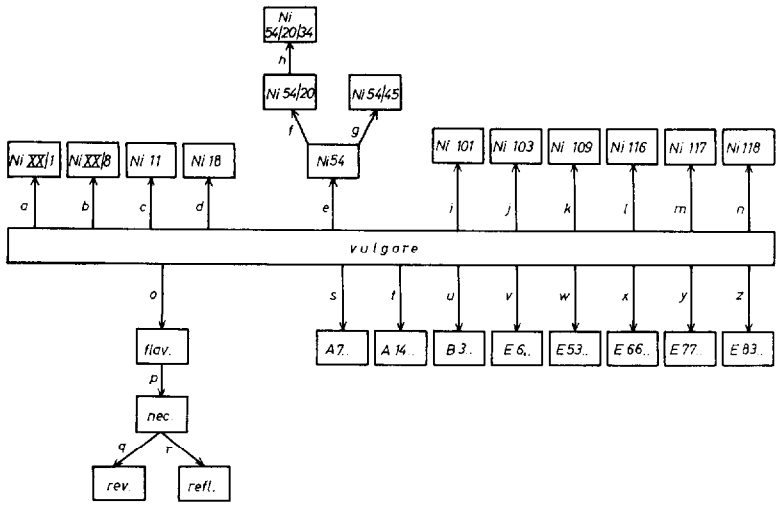

FIG. 1. Scheme of the TMV mutants studied. The differences in the amino acid composition of the tryptic peptides are given in the text.

(flav.), *necans*, (nec.), *revirescens* (rev.), and *reflavescens* (refl.)], some of which were isolated many years ago (*9*). The amino acid composition of the latter mutants has been analyzed before, without splitting the protein into tryptic peptides (*10*).

When t is the time of treatment and τ the time after which the infectivity decreases to $1/e$ of the value before treatment [see (*2*)], then for both Ni XX/1 and Ni XX/8, $\frac{t}{\tau} = 1.1$; for Ni 11, Ni 18, and Ni 54, $\frac{t}{\tau} = 5.2$ [see (*11*)]; for Ni 54/45 and Ni 54/20, $\frac{t}{\tau} = 2.6$; for Ni 54/20/34, $\frac{t}{\tau} = 6.5$; and for Ni 101 to Ni 118, $\frac{t}{\tau} = 3.0$. The 6 mutants Ni 101 to Ni 118 and the 2 Ni XX/... were produced by treatment of the virus nucleoprotein with nitrous acid whereas all other Ni mutants were isolated after nitrous acid treatment of the TMV-RNA. The mutants *flavum*, Ni XX/8, Ni 54/20/34, and E 83 are yellow strains which cannot be differentiated by symptoms on "Samsun tobacco".

The analysis of the tryptic peptides of the 26 chemically induced and spontaneous mutants listed, gives the following results [cf. Fig. 1; peptides numbered as in (*7*)]:

Steps a, e, d, e, f, g, s, u, v, w, y: no change
Steps b, h, o, z: Asp → Ala in peptide XII[1]
Steps i, j: Asp → Gly in XII
Step k: Glu → Gly in VII
Step l: Asp → Gly in IX
Step m: Ser → Phe in IV; Glu → Val in XII
Step n: Pro → Leu in XII
Step p: Ala + Phe → Val + Leu in XII; Ser → Phe in IV
Step q: Val + Leu → Asp + Phe in XII
Step r: Leu → Phe in XII
Step t: Ileu → Thr in XI
Step x: Asp → Lys in IV

The results confirm the earlier finding (*5*) that there exist mutants which have the same amino acid composition of their tryptic peptides although they show very striking differences in symptoms; this is true for both spontaneous and chemically induced mutants. This can be explained by assuming that the region within the TMV-RNA which codes for the 157 amino acids of the virus protein is smaller than the 6500 nucleotides making up the TMV-RNA. Only when nucleotides within this region are changed does an alteration of amino acids in the viral protein occur, whereas changes of nucleotides outside this region result

[1] Peptide No. XII (with 41 amino acids) contains 4 Asp [see (*7*)]. ASP means *aspartic acid or asparagine*. Glu means *glutamic acid* or glutamine.

in alterations that lead to different symptoms without changing the viral protein. By investigation of more nitrous acid mutants than have been studied hitherto, the number of nucleotides within the "gene" for the viral protein could be determined and the question answered how many nucleotides code for one amino acid. The elucidation of the total sequence of the 157 amino acids of the TMV protein (*12*) makes it possible to pinpoint the alteration of a certain amino acid at a certain position within the protein chain. This opens up the possibility of seeing whether the points of amino acid replacement are distributed randomly over the protein chain or whether there exist regions with a high frequency of replacement, and also of answering the question whether a certain amino acid can be substituted only by a limited number of others; this information would give additional help in resolving the code.

ACKNOWLEDGMENT

The technical assistance of Miss Ingrid Hindennach, Miss Brigitte Ostertag, and Miss Julita Huf is gratefully acknowledged.

REFERENCES

1. SCHUSTER, H., and SCHRAMM, G., *Z. Naturforsch.* **13b,** 697–704 (1958).
2. GIERER, A., and MUNDRY, K. W., *Nature* **182,** 1457–1458 (1958).
3. WITTMANN, H. G., *Symposium on Genetic Determination of Protein Structure, Ann. Meeting Natl. Acad. Sci. Washington April, 1960.*
4. TSUGITA, A., and FRAENKEL-CONRAT, H., *Proc. Natl. Acad. Sci. U. S.,* **46,** 636–642 (1960).
5. WITTMANN, H. G., *Z. Vererbungslehre* **90,** 463–475 (1959).
6. SPACKMAN, D. H., STEIN, W. H., and MOORE, S., *Anal. Chem.* **30,** 1130–1206 (1958).
7. WITTMANN, H. G., and BRAUNITZER, G., *Virology* **9,** 726–728 (1959).
8. MELCHERS, G., unpublished data.
9. MELCHERS, G., *Naturwissenschaften* **36,** 48–49 (1942).
10. AACH, H. G., *Z. Naturforsch.* **13b,** 425–433 (1958).
11. MUNDRY, K. W., and GIERER, A., *Z. Vererbungslehre* **89,** 614–630 (1958).
12. ANDERER, A., UHLIG, H., WEBER, E., and SCHRAMM, G., *Nature* **186,** 922–925 (1960).

Max-Planck-Institut für Biologie H. G. WITTMANN
Abt. Melchers
Tübingen, Germany
Received October 21, 1960